graphab 1.1

Manuel d'utilisation

Sommaire

1. Introduction *4*
1.1. A propos de Graphab 4
1.2. Configuration requise 4
1.3. Installation et lancement du programme 4

2. Démarrer un projet Graphab *5*
2.1. Identification du projet 5
2.2. Importation de la carte du paysage et définition des nœuds 6
2.3. Création d'un jeu de liens 7

3. Création d'un graphe *9*

4. Capacité des taches *10*
4.1. Capacité calculée comme une fonction du voisinage 10
4.2. Capacité définie à partir de données externes 11

5. Calcul des métriques de connectivité *12*
5.1. Famille de métriques et niveaux de calcul 12
5.2. Paramétrage des métriques pondérées 13
5.3. Calcul des métriques par lot 14
5.4. Interpolation des métriques 16

6. Liens entre graphes et données ponctuelles *17*
6.1. Import de données ponctuelles 17
6.2. Matrices de distance inter-points 18
6.3. Générer des points aléatoires 18
6.4. Modèle de distribution d'espèce 19

7. Affichage *21*
7.1. Propriétés des graphes 21
7.2. Propriétés des objets 22

8. Limitations *23*
8.1. Capacités de traitement et limitations 23

9. Annexes *25*
9.1. Détail de calcul des métriques 25
9.2. Références bibliographiques 32

1. Introduction

1.1. A propos de Graphab

Le programme Graphab est un outil de modélisation des réseaux écologiques fondé sur les graphes paysagers. C'est un outil intégré composé de 4 modules :

- création de graphes à partir d'une carte de paysage avec identification des taches d'habitat et des liens (distance euclidienne ou chemin de moindre coût)
- calcul d'un grand nombre de métriques de connectivité
- intégration des métriques de connectivité dans un modèle de distribution d'espèce
- interface de géo-visualisation

1.1.1. Auteurs

Le programme Graphab a été développé par Gilles Vuidel et Jean-Christophe Foltête au laboratoire ThéMA (Université de Franche-Comté – CNRS). Le développement du logiciel a été financé par le ministère de l'écologie, du développement durable, des transports et du logement dans le cadre du programme ITTECOP. Le logo de Graphab a été conçu par Gachwell.

1.1.2. Conditions d'utilisation

Le programme Graphab est disponible au téléchargement gratuitement pour un usage non-commercial. Les utilisateurs de Graphab sont invités à citer la référence suivante dans leurs travaux :

Foltête J.C., Clauzel C., Vuidel G., 2012. A software tool dedicated to the modelling of landscape networks. *Environmental Modelling & Software*, 38: 316-327.

Pour toute autre utilisation, un accord préalable du laboratoire ThéMA est nécessaire, pour cela envoyer votre demande à l'adresse : graphab@univ-fcomte.fr

1.2. Configuration requise

Graphab fonctionne sur n'importe quel ordinateur supportant Java 1.6 ou supérieur (PC sous Linux, Windows, Mac...). Toutefois, lorsqu'il s'agit de données très volumineuses, la quantité de mémoire vive (RAM) de l'ordinateur peut limiter le nombre maximum de nœuds et de liens qui peuvent être traités en une seule fois avec Graphab. De plus, pour certaines métriques complexes, la puissance de votre processeur va déterminer la vitesse de leur calcul. Vous trouverez plus de détails à ce sujet dans la rubrique 8 et dans la référence citée plus haut.

1.3. Installation et lancement du programme

Graphab est téléchargeable à cette adresse : http://thema.univ-fcomte.fr/productions/graphab.

- Télécharger et installer Java 1.6 ou + (java.com). Si vous avez un système d'exploitation 64 bits, installez de préférence la version 64 bits de Java.
- Télécharger et décompresser graphab.zip
- Lancer Graphab.jar

Une fois Graphab.jar lancé, le menu fichier donne accès à 4 rubriques :

- **Fichier / Nouveau projet** : pour commencer la création d'un projet dans lequel toutes les données et résultats sont automatiquement enregistrés.
- **Fichier / Ouvrir un projet**: ouvre un projet déjà créé auparavant.
- **Fichier / Préférences** : pour changer certains paramètres du programme : langue anglais/français ; quantité maximale de mémoire utilisable ; nombre de processeurs à utiliser. Il est recommandé d'adapter la quantité de mémoire et le nombre de processeurs aux capacités de l'ordinateur utilisé et aux traitements prévus (voir à ce sujet la rubrique 8).

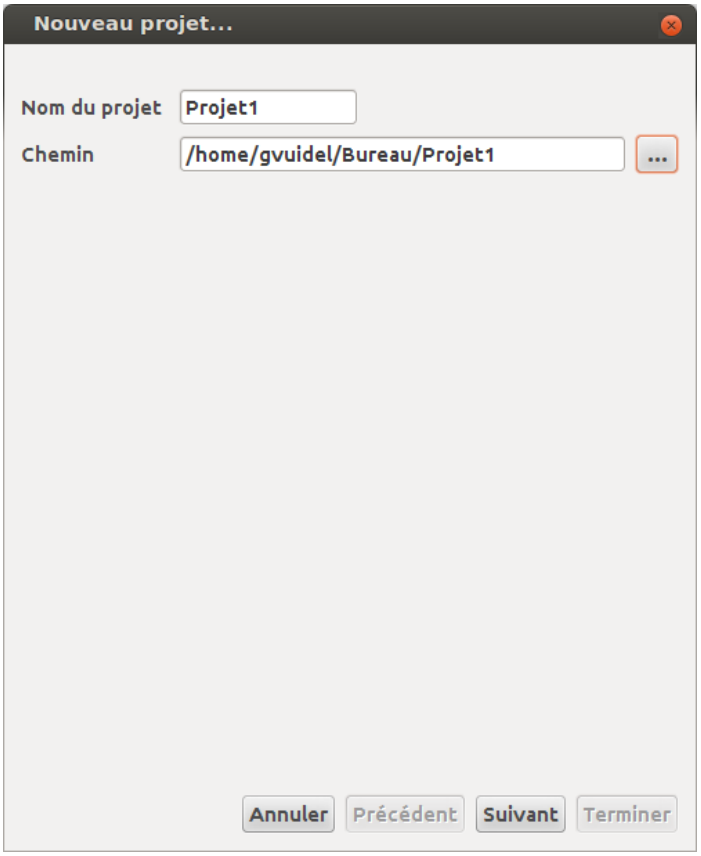
- **Fichier / Journal** : affiche le journal des évènements.

2. Démarrer un projet Graphab

La mise en place d'un nouveau projet Graphab s'effectue à partir du menu : **Fichier / Nouveau projet**. L'utilisateur doit renseigner plusieurs fenêtres successives : identification du projet, importation de la carte paysagère, création d'un jeu de liens. Un projet correspond à l'exploitation d'une seule carte paysagère initiale, mais il peut comporter plusieurs jeux de liens. Au terme de la phase de démarrage, le projet sera le support pour la création de multiples graphes et calculs de métriques de connectivité.

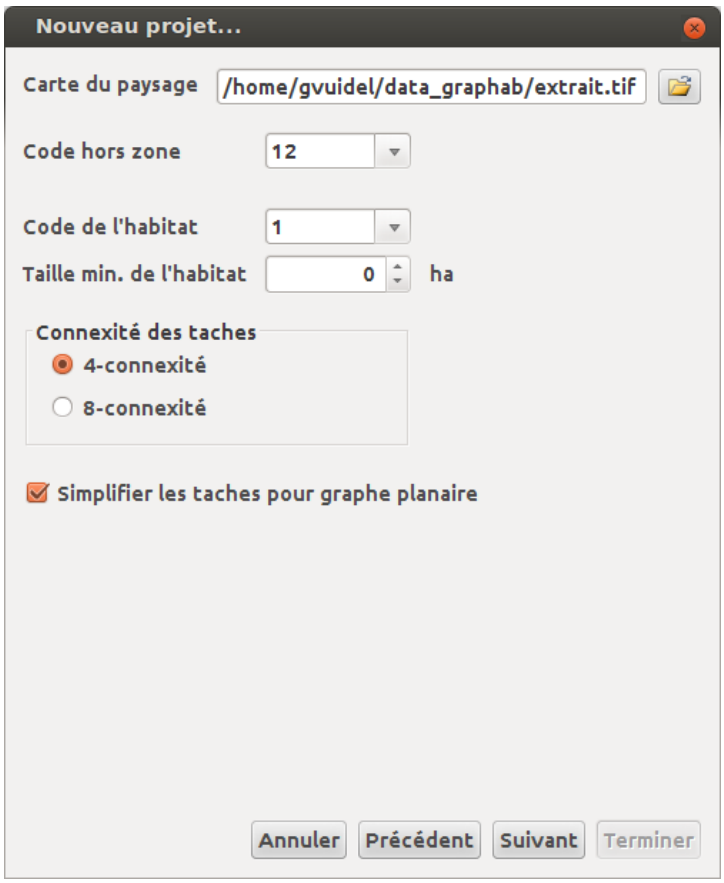
2.1. Identification du projet

Dans la première fenêtre, l'utilisateur doit nommer le projet et indiquer dans quel dossier celui-ci sera créé.



2.2. Importation de la carte du paysage et définition des nœuds

La deuxième fenêtre concerne l'importation de la carte du paysage. Celle-ci doit être un fichier de type raster (*.tif, *.rst) dans lequel la valeur de chaque pixel correspond à une catégorie (occupation du sol ou autre type de classification).



S'il est en format *.tif, sans extension Geotiff, le fichier doit être associé à un fichier de géoréférencement (*.tfw) structuré comme suit :

Exemple	
10	Taille du pixel en X
0	Rotation sur les lignes (0 pour les bonnes images)
0	Rotation sur les colonnes (0 pour les bonnes images)
-10	Taille du pixel en Y
821755	Coordonnée en X du pixel haut gauche
2342995	Coordonnée en Y du pixel haut gauche

S'il est en format *.rst, le fichier doit être associé à un fichier de géoréférencement (*.rdc) suivant la structure proposée dans le logiciel Idrisi.

Important ! L'unité du système de coordonnées de l'image doit être en mètre. Si ce n'est pas le cas, les unités de surface et de distance seront erronées. Il est possible de reprojeter l'image dans une projection métrique (Lambert93, UTM, …) à l'aide d'un SIG.

Code hors zone : code des pixels correspondant à l'absence de valeurs dans le fichier raster.

Code de l'habitat : valeur des pixels correspondant à l'habitat utilisé pour la définition des taches.

Taille min. de l'habitat : surface minimale en hectare pour qu'une tache d'habitat devienne un nœud du graphe.

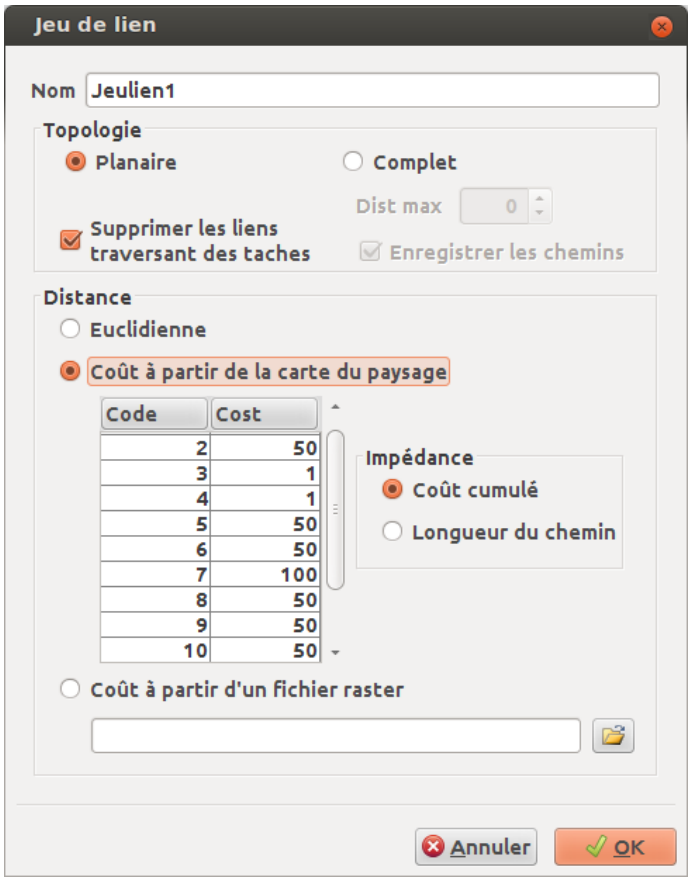
Connexité des taches :

- 4-connexité : une tache est constituée du pixel central et de ses 4 voisins s'ils ont la même valeur ;
- 8-connexité : une tache est constituée du pixel central et de ses 8 voisins s'ils ont la même valeur.

Simplifier les taches pour graphe planaire : cette option, lorsqu'elle est cochée, permet d'accélérer la création d'un graphe planaire, en simplifiant les limites des polygones des taches. Ce processus de simplification n'étant pas déterministe, la création de deux graphes planaires pour une même carte du paysage peut engendrer des limites de polygones légèrement différentes. Par conséquent, cette option ne doit pas être cochée dans le cas d'une comparaison de graphes planaires.

2.3. Création d'un jeu de liens

La troisième fenêtre concerne la création d'un jeu de liens, ce qui implique la sélection de plusieurs paramètres : topologie et pondération des liens. La création de ce jeu de liens constitue la dernière étape dans la mise en place initiale du projet. Cependant, il est possible de créer à tout moment de nouveaux jeux de liens dans le même projet, par le menu : **Graphes / Créer un jeu de liens**.



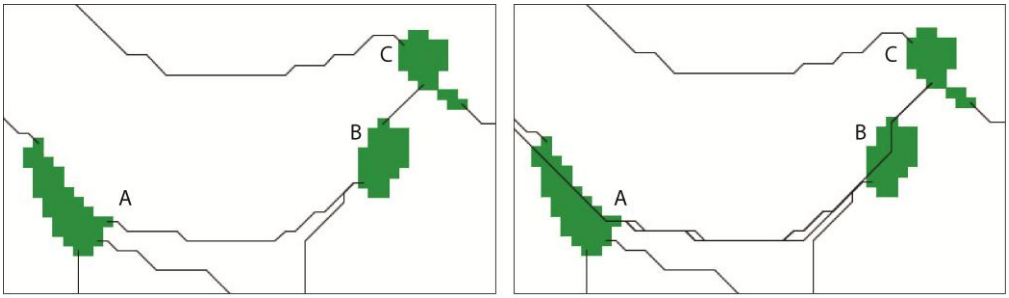
2.3.1. Topologie des liens

Deux topologies sont disponibles :

- planaire : seuls les liens formant un graphe planaire minimal sont pris en compte. Cette topologie est mise en place par l'intermédiaire de polygones de Voronoï autour de chaque tache d'habitat. Ces polygones sont définis à partir des bords des taches, en distance euclidienne.
- complet : tous les liens entre les taches sont potentiellement pris en compte.

Distance max : Cette option disponible pour la topologie complète permet de préciser un seuil de distance au-delà duquel les liens ne sont plus créés. Elle permet ainsi de limiter le nombre de liens créés et d'accélérer la création des liens en topologie complète.

Supprimer les liens traversant les taches : cette option cochée par défaut permet de ne pas créer de lien entre deux taches (A et C sur la figure suivante) passant par une tache intermédiaire (B). Ceci est recommandé pour le calcul de la métrique de centralité intermédiaire (BC) pour prendre en compte la fréquence avec laquelle chaque tache se trouve sur les plus courts chemins liant les paires de nœuds sur le graphe. A noter : cette option fournit une approximation de la distance intra-taches entre A et C. Si la case est décochée, un lien entre A et C passant par B sera créé représentant la distance complète réelle entre les deux taches.



Option cochée : les liens traversant les taches sont supprimés

Option décochée : les liens traversant les taches sont conservés

Illustration de l'option « Supprimer les liens traversant les taches »

Enregistrer les chemins :

- si la case est cochée : les liens sont enregistrés sous forme de chemins représentant le tracé réel du lien entre deux taches.
- si la case est décochée : les liens sont enregistrés sous forme topologique uniquement. Dans ce cas, la visualisation des liens en vue réaliste n'est plus possible. Ceci est recommandé lorsque le graphe contient de nombreux liens (dans le cas d'un graphe complet non seuillé par exemple) afin de limiter l'utilisation de la mémoire. A noter : si les chemins ne sont pas enregistrés, il n'est pas possible d'inclure les distances intra-taches dans le calcul des métriques.

2.3.2. Distance (ou impédance des liens)

Les distances sont calculées de bord à bord entre les taches. Deux principaux types de distance sont possibles : distance euclidienne et distance de moindre coût.

- Distance euclidienne : les liens sont calculés en distance euclidienne (parcours à vol d'oiseau entre les taches), ce qui revient à considérer la matrice comme uniforme.

- Distance de moindre coût : les liens sont calculés en distance coût. L'hétérogénéité de la matrice est prise en compte en assignant des valeurs de résistance (ou friction) aux classes paysagères. L'utilisateur peut activer cette possibilité de deux façons :

1) soit en indiquant les coûts correspondant aux catégories de la carte du paysage dans le tableau ;
2) soit à partir d'un fichier raster externe (format *.tif ou *.rst) contenant pour chaque pixel une valeur de résistance.

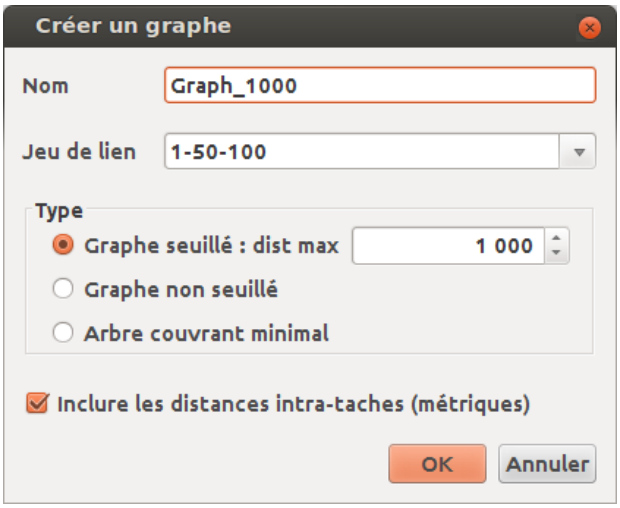
L'usage des distances de moindre coût donne accès à deux types d'impédance calculés à partir des mêmes chemins :

- coût cumulé : l'impédance est égale à la somme des coûts de tous les pixels du chemin parcouru.
- longueur du chemin de moindre coût : l'impédance est égale à la longueur métrique du chemin parcouru.

Pour chaque lien créé, sa distance métrique et sa distance en unité de coût sont enregistrées et consultables dans les propriétés (cf. partie 6).

3. Création d'un graphe

Un projet Graphab peut donner lieu à la création de plusieurs graphes. Chaque graphe étant construit à partir d'un jeu de lien, soit on utilise le jeu de lien défini lors de la mise en place initiale du projet, soit on définit un nouveau jeu de lien par le menu : **Graphes / Créer un jeu de liens**.
La création d'un graphe s'effectue par le menu : **Graphe / Créer un graphe**.

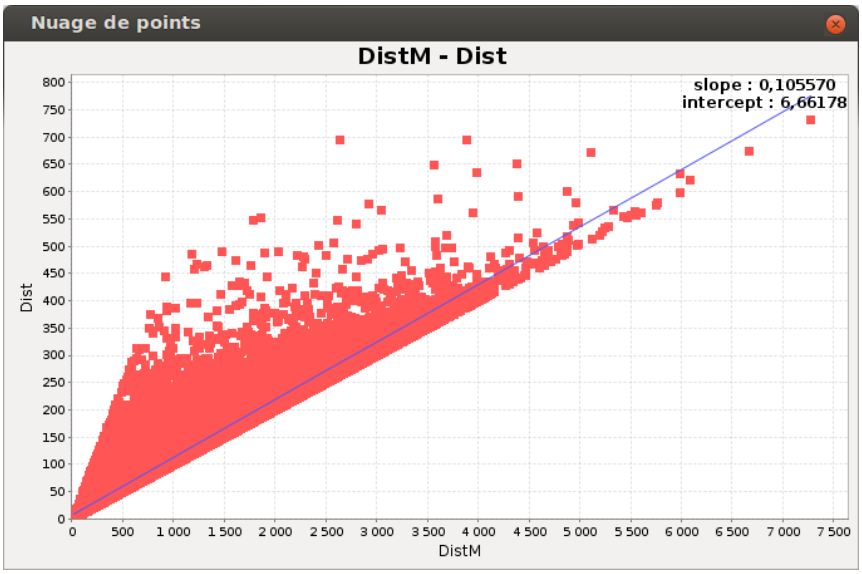


Le graphe à créer doit tout d'abord être nommé.

L'utilisateur doit ensuite sélectionner un des jeux de liens créés à l'étape 2.2, puis choisir un type de graphe :

- graphe seuillé : les liens retenus sont inférieurs ou égaux au seuil de distance choisi.
- graphe non seuillé : tous les liens sont créés entre les taches quelle que soit leur distance.
- arbre couvrant minimal : graphe reliant toutes les taches dont le poids total des liens est minimal.

Pour un graphe seuillé, l'unité de la distance de seuillage dépend du type de distance utilisé dans la création du jeu de liens. Si le jeu de lien a été créé en distance euclidienne, la distance de seuillage du graphe est exprimée en mètres. Si la création du jeu de lien a été réalisée en distance-coût, la distance de seuillage du graphe est exprimée en coûts cumulés.



Il est possible d'obtenir une approximation de la distance métrique (appelée DistM) exprimée en coûts cumulés (appelée Dist) en affichant le nuage de points correspondant au jeu de lien utilisé (cf. 6), et en utilisant la droite de la régression [Dist = intercept + slope × DistM] pour effectuer cette conversion.

Option « Inclure les distances intra-taches » : si la case est cochée, le calcul des métriques à l'étape 3 tient compte des distances intra-taches (recommandé) ; dans le cas contraire, seules les distances inter-taches sont prises en compte.

Pour réaliser une analyse multiscalaire, il est souvent nécessaire de créer une série de graphes seuillés dont le seuil est défini de façon croissante. L'utilisateur peut créer manuellement cette série de graphes, mais si l'objectif est d'analyser le comportement d'une métrique de connectivité en fonction du seuil, il est préférable d'utiliser le menu : **Métriques / Graphes par lot** (cf. 5.3).

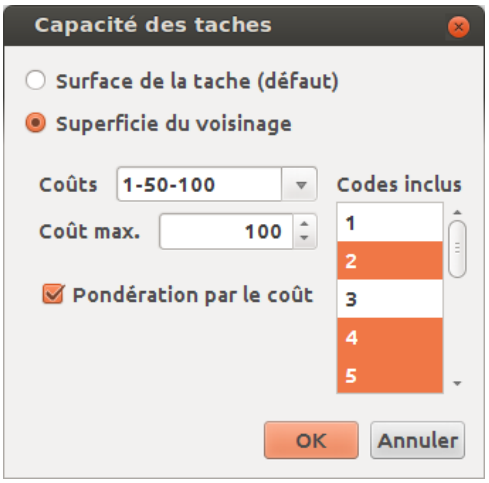
4. Capacité des taches

La capacité d'une tache traduit sa « qualité intrinsèque », considérée comme un indicateur de son potentiel démographique. Ainsi une tache avec une forte capacité pourra accueillir une forte population et inversement. La capacité intervient directement dans le calcul de certaines métriques de connectivité « de surface » et « pondérées » (voir la section 5).

Par défaut (lors de la création initiale du projet), la capacité d'une tache est égale à sa superficie en m^2. Cependant, il est possible de remplacer la superficie par n'importe quel autre indicateur de qualité. En effet, dans certains cas, la présence d'une espèce n'est pas liée à la superficie de la tache mais à la superficie d'autres types d'occupation du sol autour de la tache. Par exemple la présence d'amphibiens dans un étang de reproduction ne dépend souvent pas de la taille du plan d'eau mais de la superficie de l'habitat terrestre autour de ce plan d'eau.

4.1. Capacité calculée comme une fonction du voisinage

Avec le menu **Données / Calculer la capacité des taches**, il est possible de définir la capacité des taches comme une fonction de la composition de leur voisinage et de procéder à ce calcul directement dans Graphab.



L'utilisateur doit sélectionner plusieurs critères : type de distance, distance maximale, catégories paysagères.

Coûts : il s'agit de la métrique spatiale (euclidienne ou ensemble de coûts) correspondant à un des jeux de liens disponibles dans le projet. Il est à noter que l'usage des coûts dans cette procédure revient à définir un voisinage anisotropique autour des taches, ce qui peut différer fortement d'une fonction de type buffer. Pour des raisons de cohérence, il est conseillé d'utiliser le même type de distance que celui qui a permis de définir les liens du graphe. Dans le cas d'un jeu de lien créé avec une distance euclidienne, il faut choisir « all cost = 1 ».

Coût max. : Comme pour la distance de seuillage du graphe, l'unité de cette distance maximale de voisinage dépend du type de distance utilisé dans la création du jeu de liens (euclidien ou distance-coût).

Codes inclus : l'utilisateur peut choisir une ou plusieurs catégories paysagères, autre que la catégorie « habitat », à prendre en compte dans le calcul de la capacité.

L'option « pondération par la distance » permet d'introduire une pondération en fonction de l'éloignement à la tache, par le biais d'une fonction exponentielle négative. De cette façon, les surfaces sélectionnées compteront plus si elles sont proches de la tache et inversement.

Les valeurs de capacité ainsi calculées remplacent les surfaces des taches pour tous les calculs qui suivront, mais il est possible de revenir à l'état initial par le menu : **Données / Calculer la capacité des taches** et en choisissant la surface de la tache.

4.2. Capacité définie à partir de données externes

En activant le menu **Données / Importer la capacité des taches**, on peut ouvrir un tableau de données au format *.csv, décrivant toutes les taches du projet et contenant comme attribut les valeurs de capacité définies au préalable par l'utilisateur. Les identifiants du tableau importé doivent être ceux qui ont été définis lors de la création initiale du projet.

Les valeurs de capacité du tableau importé remplacent les surfaces des taches pour tous les calculs qui suivront cette importation, mais il est possible de revenir à l'état initial par le menu : **Données / Calculer la capacité des taches** et en choisissant la surface des taches.

5. Calcul des métriques de connectivité

5.1. Famille de métriques et niveaux de calcul

Chaque graphe créé dans un projet peut être utilisé pour calculer différentes métriques de connectivité, dont le détail de calcul et les références bibliographiques sont mentionnés en annexe. Ces calculs s'effectuent à plusieurs niveaux correspondant aux grandes entrées du menu **Métriques** (tableau 1) :

- **Métriques /Métrique globale** : les métriques caractérisent le graphe entier.
- **Métriques / Métrique par composante** : les métriques caractérisent la connectivité interne de chaque composante (ou sous-graphe).
- **Métriques / Métrique locale** : les métriques caractérisent la connectivité de chaque élément du graphe, nœud ou lien.
- **Métriques / Delta-métriques** : les métriques caractérisent également chaque élément du graphe, mais à partir d'un mode de calcul spécifique. En utilisant la méthode dite de la « suppression » (suppression de nœud ou suppression de lien), l'importance relative de chaque élément du graphe est évaluée en calculant le taux de variation que sa suppression occasionne sur une métrique globale. Le résultat d'une delta-métrique est donc local, mais en référence au niveau global.

Dans la fenêtre de dialogue qui s'ouvre suite au choix d'un de ces quatre modes de calcul, trois familles de métriques sont disponibles :

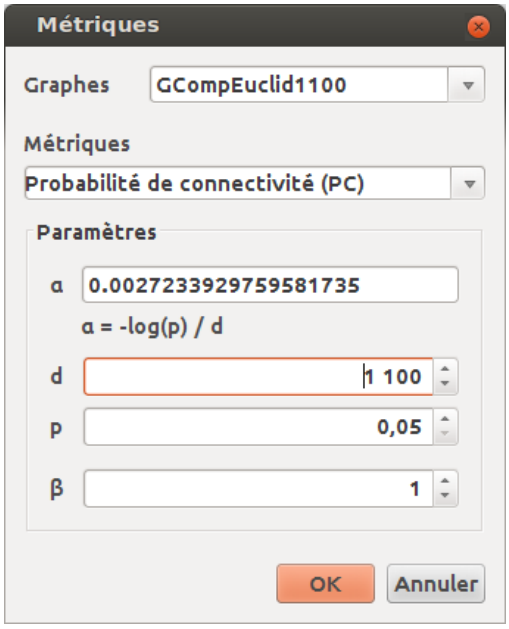
- *les métriques pondérées*, qui sont fondées sur des critères de distance et de capacité des taches et qui nécessitent un paramétrage adapté à l'espèce de référence. Ces métriques nécessitent des calculs de chemin sur les graphes, par le biais de l'algorithme de Dijkstra. En choisissant ces métriques, l'utilisateur doit indiquer le réglage souhaité.
- *les métriques de surface*, qui sont fondées en priorité sur un critère de surface. Dans le cas où la capacité ne correspond pas à la surface des taches mais à un autre critère, ces métriques peuvent être calculées et elles s'expriment dans l'unité du critère considéré.
- *les métriques topologiques*, qui sont issues de la théorie des graphes et qui ne nécessitent pas de paramétrage.

Famille	Métrique de connectivité	Code	Niveau de calcul			Delta métriques
			Global	Composante	Local	
Métriques pondérées	Flux	F	×	×	×	
	Probabilité de connectivité	PC	×	×		×
	Probabilité de connectivité flux	FPC			×	
	Fractions de delta Probabilité de connectivité	dPC				×
	Indice de centralité intermédiaire	BC			×	
	Indice intégral de connectivité	IIC	×	×		×
Métriques de surface	Taille moyenne des composantes	MSC	×			
	Taille de la plus grande composante	SLC	×			
	Probabilité de coïncidence de classe	CCP	×			
	Taille de composante attendue	ECS	×			
Métriques topologiques	Degré	Dg			×	
	Coefficient de groupement	CC			×	
	Centralité de proximité	CCe			×	
	Excentricité	Ec			×	
	Corrélation de connectivité	CCor			×	
	Elément de coupure	Cut			×	
	Nombre de composantes	NC	×			
	Diamètre	GD	×	×		×
	Indice d'Harary	H	×	×		×

Tableau 1. Liste des métriques de connectivité et de leur niveau de calcul. Les codes sont indiqués suivant le nom anglais des métriques.

Quel que soit le niveau choisi, l'utilisateur doit d'abord indiquer le graphe sur lequel le calcul va s'appliquer, puis choisir la métrique de connectivité.

5.2. Paramétrage des métriques pondérées



5.2.1. Paramètre alpha

Plusieurs métriques intègrent dans leur calcul une pondération où les distances entre les taches sont converties en probabilité de déplacement. Ces métriques sont : F, FPC, PC, BC. La pondération est basée sur une fonction exponentielle de la forme :

$$p = e^{-\alpha d}$$

où *p* est la probabilité de déplacement entre deux taches, *d* la distance entre ces taches et α un paramètre contrôlant la vitesse avec laquelle la probabilité diminue quand la distance augmente. Comme la valeur du paramètre α n'est pas facile à déterminer, le logiciel propose de le calculer à partir des deux autres paramètres. L'utilisateur doit indiquer la distance correspondant à une certaine valeur de probabilité, par exemple :

- la distance de dispersion maximale de l'espèce correspondant à une faible valeur de *p* (0.05 ou 0.01).
- la distance de dispersion moyenne de l'espèce correspondant à une valeur médiane de *p* (0. 5).

La valeur α est automatiquement obtenue à partir de la formule :

$$\alpha = -log(p)/d$$

Dans le cas d'un graphe seuillé, il est supposé que la distance *d* utilisée dans ce paramétrage soit en cohérence avec celle qui a été utilisée pour le seuillage du graphe.

5.2.2. Paramètre béta

Les mêmes métriques F, PC, FPC, BC sont contrôlées par le paramètre β. Ce paramètre est l'exposant appliqué à la capacité des taches. Il joue sur l'équilibre relatif entre le poids des distances et le poids des capacités des taches dans la pondération des métriques. En prenant l'exemple de la métrique F en calcul local, dont la forme générique est la suivante :

$$F = \sum a^{\beta}\, e^{-\alpha d}$$

- une valeur de $\beta = 0$ signifie que la capacité des taches ne joue pas de rôle dans la pondération.

- une valeur de $\beta = 1$ signifie que la capacité des taches joue de façon linéaire dans cette pondération.

- une valeur de $\beta = 2$ signifie que la capacité des taches est portée à la puissance 2.

- une valeur de $\beta = 0.5$ signifie que la racine carrée de la capacité des taches intervient dans la pondération.

- une valeur de $\beta = -1$ signifie que la capacité des taches joue de façon inversement proportionnelle.

Au-delà de ces quelques exemples, toutes les valeurs de pondération sont possibles.

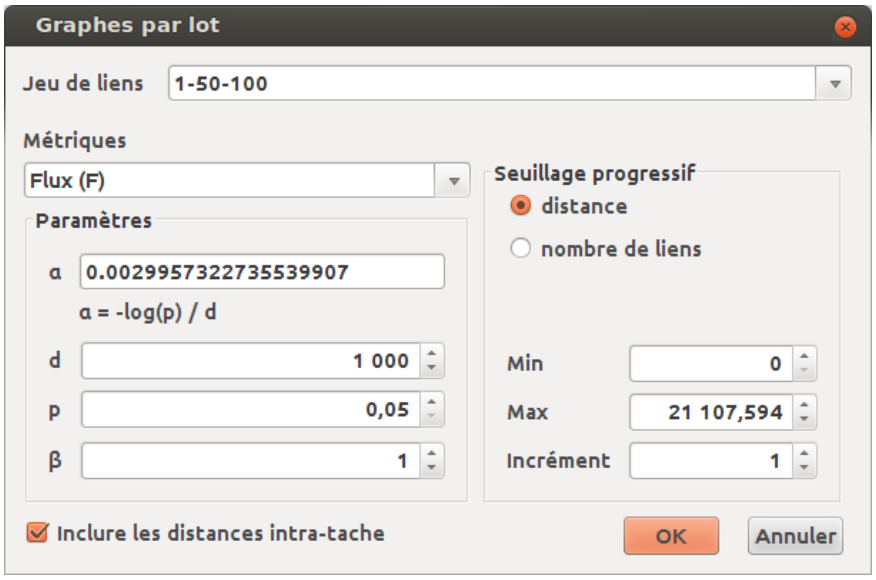
5.3. Calcul des métriques par lot

Chaque métrique compatible avec le niveau global peut être calculée en chaîne suivant la variation de l'échelle des distances. Cette variation peut concerner soit le seuillage du graphe (5.3.1.), soit le paramétrage des métriques elles-mêmes (5.3.2.). Le type de distance considéré pour le seuillage est celui qui a été défini pour le jeu de lien choisi, à savoir distance euclidienne, distance de moindre coût, chemin de moindre coût.

5.3.1. Graphes par lot

Le menu **Métriques / Graphes par lot** permet de créer une série de graphes seuillés à partir d'un jeu de liens donné, et de calculer pour chacun une métrique de niveau global. Les seuils des graphes successifs sont définis de façon croissante, suivant deux modes :

- distance : un intervalle de distance fixe est déterminé entre les graphes successifs. Les valeurs de la métrique sont donc calculées à des intervalles réguliers de distance.
- nombre de liens : un nombre de liens fixe est défini entre les graphes successifs. Ce nombre de liens est automatiquement transformé en distance utilisée pour seuiller le graphe. Ces distances seuil peuvent s'échelonner de façon irrégulière.



Les graphes sont définis suivant trois critères donnés par l'utilisateur :

- minimum : plus petit seuil, utilisé pour le premier graphe de la série. Par défaut, ce minimum est à 0, ce qui correspond à l'absence totale de liens.
- maximum : seuil maximal, utilisé pour le dernier graphe de la série. Par défaut, ce maximum correspond à la distance maximale, ou le nombre de lien, du jeu de lien sélectionné.
- incrément : valeur de distance ajoutée à chaque nouveau graphe.

Une fois le calcul terminé, le logiciel ouvre une nouvelle fenêtre dans laquelle la courbe de la métrique choisie est affichée en fonction du seuil de distance. Les valeurs de cette courbe peuvent être enregistrées par le bouton export, en sélectionnant le format texte.

5.3.2. Paramètres par lot

Le menu **Métriques / Paramètres par lot** permet de calculer en série des métriques à partir d'un graphe donné. Cette possibilité s'applique exclusivement aux métriques pondérées. Elle se subdivise en deux entrées : métriques locales ou globales.

Paramètres par lot pour métriques locales

Une métrique pondérée de niveau local est calculée en série suivant la variation d'un de ses paramètres. L'utilisateur doit choisir le graphe, la métrique et le paramètre à faire varier. La variation du calcul est déterminée par :

- min : valeur minimale du paramètre
- max : valeur maximale du paramètre
- incrément : valeur d'intervalle entre deux calculs de la métrique

Une fois le calcul terminé, les taches (et dans certains cas les liens) du graphe sont caractérisées par une série de métriques supplémentaires.

Paramètres par lot pour métriques globales

Pour un graphe donné, une métrique pondérée de niveau global est calculée en série suivant la variation d'un de ses paramètres. Comme précédemment, cette variation est définie entre une valeur minimale (min), une valeur maximale (max) et un intervalle (incrément).

La procédure se termine par l'ouverture d'une nouvelle fenêtre dans laquelle la courbe de la métrique choisie est affichée en fonction du paramètre.

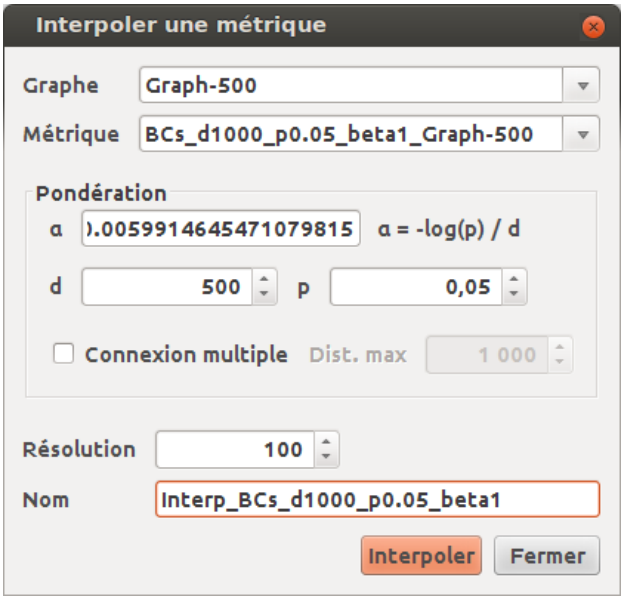
Le tableau 2 récapitule les possibilités de calcul des métriques.

Type	Métrique de connectivité	Code	Capacité des taches	Distance intra-taches	Paramètres		Graphe par lot	Paramètre par lot
					α	β		
Métriques pondérées	Flux	F	×	×	×	×	×	×
	Probabilité de connectivité	PC	×	×	×	×	×	×
	Probabilité de connectivité flux	FPC	×	×	×	×		
	Fractions de delta Probabilité de connectivité	dPC	×	×	×	×		
	Centralité intermédiaire	BC	×	×	×	×		×
	Indice Intégral de Connectivité	IIC	×				×	
Métriques de surface	Taille moyenne des composantes	MSC	×				×	
	Taille de la plus grande composante	SLC	×				×	
	Probabilité de coïncidence de classe	CCP	×				×	
	Taille de composante attendue	ECS	×				×	
Métriques topologiques	Degré	Dg						
	Coefficient de groupement	CC						
	Centralité de proximité	CCe		x				
	Excentricité	Ec		x				
	Corrélation de connectivité	CCor						
	Nombre de composantes	NC					×	
	Diamètre	GD		×			×	
	Indice d'Harary	H					×	

Tableau 2. Possibilités de calcul des métriques de connectivité.

5.4. Interpolation des métriques

Le menu **Analyse / Interpolation des métriques** permet de générer des couches raster à partir des métriques locales calculées au niveau des taches. Cette transformation est fondée sur une interpolation spatiale spécifique, qui permet d'attribuer les valeurs de connectivité des taches à chaque cellule d'une grille, en utilisant une fonction de pondération décroissante à partir de la bordure des taches (poids de 1). Globalement, plus les cellules sont situées à l'écart du graphe, plus elles obtiennent des valeurs faibles de connectivité.



La pondération est une fonction exponentielle négative de la forme $p = e^{-\alpha d}$ pour laquelle l'utilisateur doit choisir une distance (*d*) correspondant à une certaine probabilité (*p*) et le logiciel en déduit la valeur du paramètre α. En principe, ce réglage est à rendre cohérent avec le choix du graphe de référence ou le choix des éventuelles métriques pondérées qui sont à mobiliser, en utilisant la même valeur de *d*.

L'option **Connexion multiple** permet de faire intervenir plusieurs taches dans le calcul des métriques au niveau des points. Le calcul est fondé sur une moyenne pondérée des valeurs issues de toutes les taches comprises dans le voisinage des points, jusqu'à la **Distance maximale** indiquée.

Il est à noter que la distance considérée dans ces calculs dépend du graphe de référence. Si celui-ci est fondé sur une distance de moindre coût, l'interpolation spatiale utilise cette même distance et non une simple distance euclidienne.

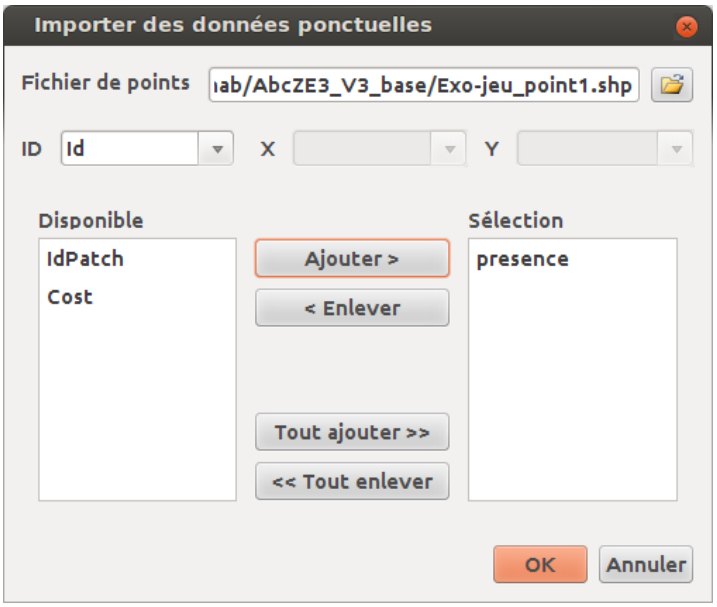
L'interpolation des métriques est aussi utilisée automatiquement dans le calcul des modèles de distribution d'espèce : voir **6.3. Modèles de distribution d'espèce**.

6. Liens entre graphes et données ponctuelles

Le corps principal du logiciel est la mise en place des graphes et le calcul de métriques de connectivité, mais il est souvent très utile de relier ces éléments à des données externes. Le logiciel permet de faire interagir les données issues des graphes avec un jeu de données de points.

6.1. Import de données ponctuelles

Des données ponctuelles peuvent être importées par le menu : **Données / Importer des données ponctuelles**. Ces données peuvent contenir plusieurs attributs, mais seuls les attributs binaires (présence/absence) seront pris en compte dans certaines procédures (voir Modèles de distribution d'espèce au 6.3).



Le fichier importé peut être soit sous forme de couche vectorielle au format shapefile (*.shp), soit sous forme de tableau (*.csv). Dans le second cas, il faut indiquer les colonnes du tableau qui correspondent aux

identifiants et aux coordonnées en X et en Y des points. Les attributs à prendre en compte doivent être sélectionnés à partir de la liste des attributs disponibles.

Si les données ponctuelles ne contiennent pas d'attribut d'absence, elles ne peuvent pas être utilisées dans des modèles de distribution d'espèce. Si l'utilisateur envisage de mettre en place un modèle de distribution d'espèce, l'utilisateur doit créer un jeu de données de pseudo-absence avec **Données / Générer des points aléatoires**.

6.2. Matrices de distance inter-points

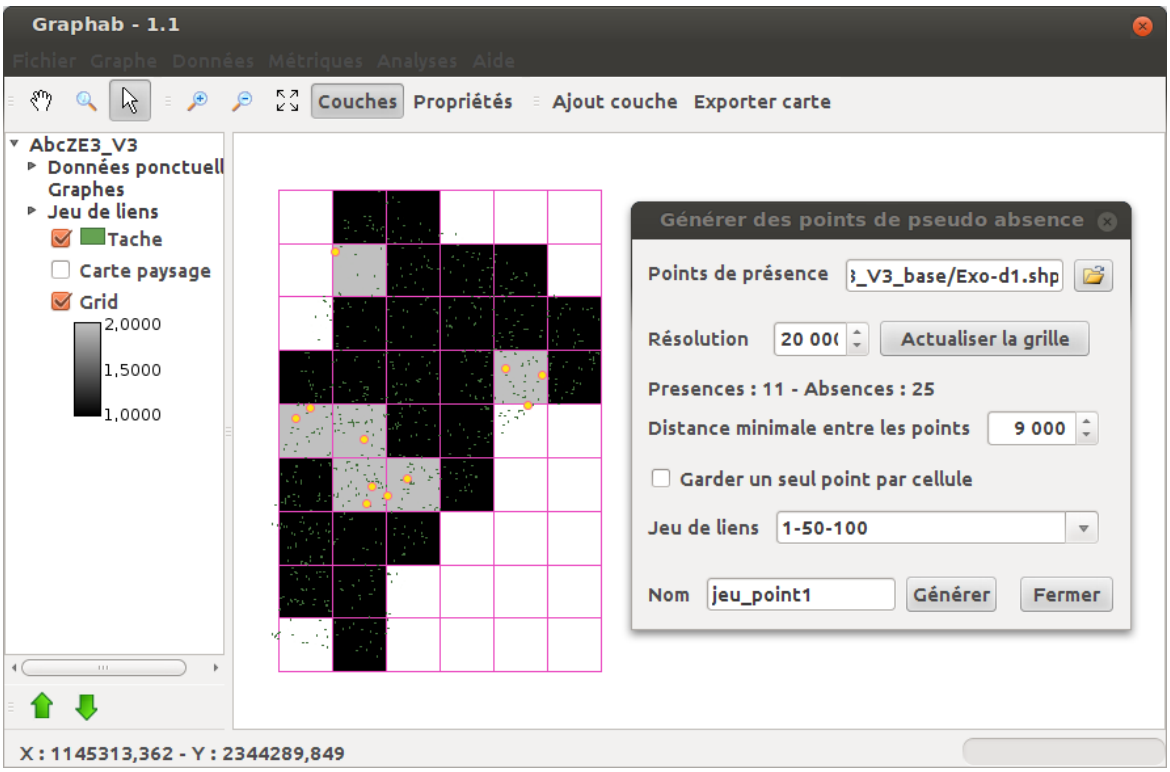
Les données ponctuelles importées dans Graphab peuvent être à l'origine du calcul de matrices de distance inter-points par un clic droit sur le nom d'un jeu de données. Plusieurs types de distances sont disponibles :

- distance raster : distance calculée en mode raster en fonction d'un jeu de lien de référence. Selon le jeu de lien choisi, il peut s'agir de la distance euclidienne, de la distance coût cumulée ou de la longueur du chemin de moindre coût. Dans les deux derniers cas, le calcul intègre les coûts associés aux catégories de la carte de paysage, tels qu'ils ont été définis dans le jeu de liens. Le résultat est une matrice de distance indépendante du graphe ; cette matrice correspond au calcul offert par les Systèmes d'Information Géographique.

- distance de graphe : distance calculée suivant le plus court chemin sur un graphe de référence à indiquer. Le type de distance correspond à celui qui a été utilisé pour la définition du jeu de lien qui a servi à définir le graphe de référence. Aux deux extrémités du chemin considéré, le calcul intègre la distance entre chaque point et la tache la plus proche. Suivant le choix qui a été fait au moment de la création du graphe, le calcul intègre éventuellement les distances intra-taches.

6.3. Générer des points aléatoires

Le menu **Données / Générer des points aléatoires** permet de créer un ensemble de points de pseudo-absence en s'appuyant sur un ensemble de points de présence.



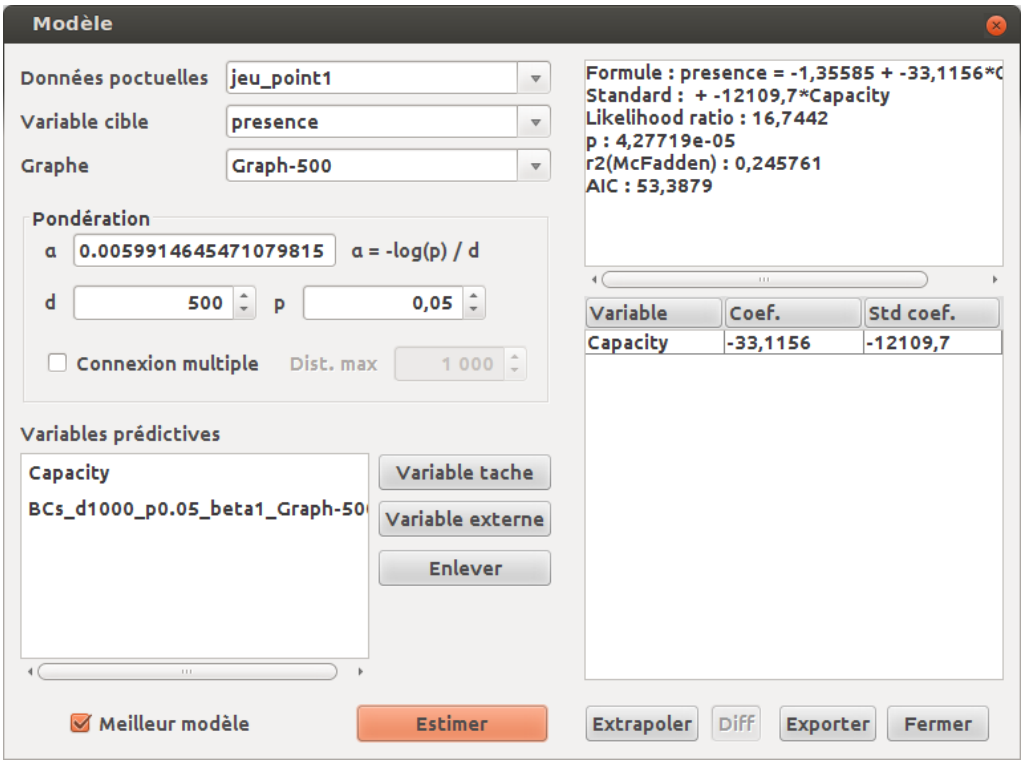
L'utilisateur doit charger un fichier de points de présence et indiquer le nom du jeu de points de présence / absence à créer.

Plusieurs paramètres sont à définir pour effectuer un tirage aléatoire des points d'absence :

- Résolution de la grille (en mètres) pour définir la taille des cellules dans lesquelles seront éventuellement tirés des points d'absence. Le bouton « actualiser la grille » permet d'afficher la grille selon la résolution choisie
- Distance minimale entre les points (en mètres) : cette fonction permet d'atténuer les effets d'autocorrélation spatiale en indiquant une distance minimale à respecter entre les points d'absence générés et entre ces points et les points de présence.
- Type de distance : l'unité de la distance minimale entre les points dépend du type de distance utilisé dans le jeu de liens sélectionné.
- Garder un point par cellule : cette option permet de ne garder qu'un point de présence dans chaque cellule pour atténuer les effets d'autocorrélation spatiale.

6.4. Modèle de distribution d'espèce

Si un jeu de points de présence / absence a été défini, le logiciel permet d'utiliser les métriques de connectivité calculées à partir d'un graphe comme variables prédictives dans un modèle de distribution d'espèce, suivant le menu **Analyse / Modèle de distribution d'espèce**. Cette modélisation est possible même si les points ne sont pas situés dans les taches d'habitat, par le biais d'une extrapolation spatiale des valeurs des métriques. Le modèle est une régression logistique fondé sur la minimisation du critère AIC.



Au préalable, l'utilisateur doit indiquer :
- le jeu de données ponctuelles à analyser ;
- la variable cible dans le modèle prédictif ;
- le graphe de référence pour l'utilisation des métriques de connectivité.

6.4.1. Pondérations pour l'extrapolation des métriques au niveau des points

Les valeurs des métriques sont calculées pour chaque point d'analyse par le biais d'une interpolation spatiale. Celle-ci est fondée sur une pondération des valeurs par une fonction décroissante à partir du bord des taches (poids de 1). Le poids décroit selon une fonction exponentielle négative de la même forme que celle qui est utilisée pour les métriques pondérées (voir 5.2.1), c'est pourquoi on retrouve le même réglage.

L'utilisateur choisi une distance (*d*) correspondant à une certaine probabilité (*p*) et le logiciel en déduit la valeur du paramètre α. En principe, ce réglage est à rendre cohérent avec le choix du graphe de référence ou le choix des éventuelles métriques pondérées qui sont à mobiliser, en utilisant la même valeur de *d*.

L'option **Connexion multiple** permet de faire intervenir plusieurs taches dans le calcul des métriques au niveau des points. Le calcul est fondé sur une moyenne pondérée des valeurs issues de toutes les taches comprises dans le voisinage des points, jusqu'à la **Distance maximale** indiquée.

Les détails de cette forme de pondération sont donnés dans Foltête et al., 2012a.

6.4.2. Estimation du modèle

Le choix d'un graphe pour appliquer le modèle fait apparaître toutes les métriques de connectivité disponibles parmi les variables prédictives.

Des métriques dérivées d'autres graphes peuvent être ajoutées en cliquant sur le bouton **Variable tache**.

Des variables provenant d'autres sources peuvent aussi être ajoutées, en cliquant sur le bouton **Variable externe**.

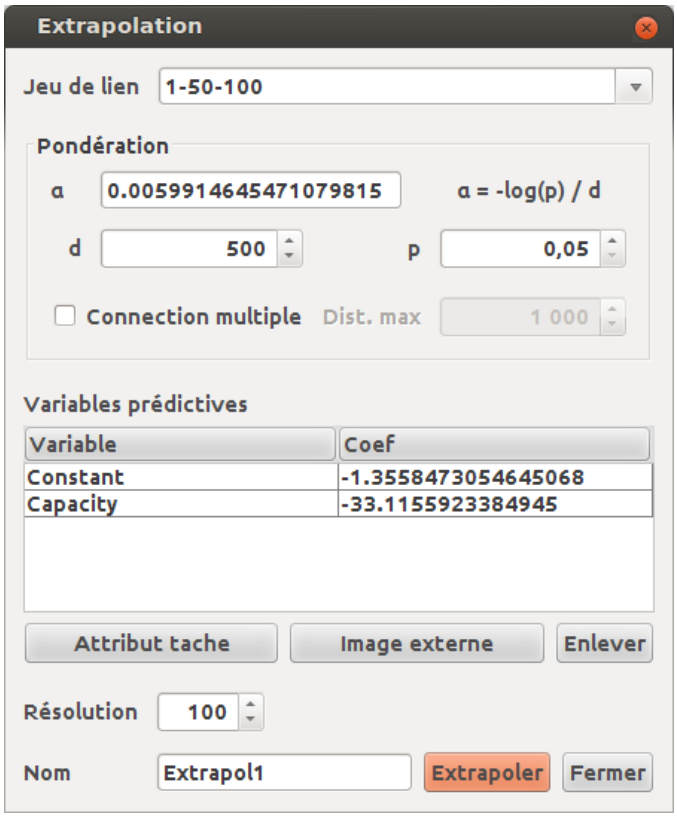
Une fois que les variables prédictives que l'utilisateur veut intégrer sont en place, le bouton **Estimer** permet de calculer les coefficients de la régression logistique. Les résultats s'affichent sur la partie droite de la fenêtre. L'option **Meilleur modèle** consiste à tester toutes les combinaisons de variables et de sélectionner celle qui minimise le critère AIC.

6.4.3. Utilisation du modèle

Un modèle prédictif considéré comme valide peut être utilisé de plusieurs façons.

Suivant le bouton **Exporter**, le tableau contenant toutes les variables statistiques impliquées dans la régression peut être exporté au format *.csv.

Le bouton **Extrapoler** permet d'estimer la probabilité de présence de l'espèce dans toutes les cellules d'une grille.



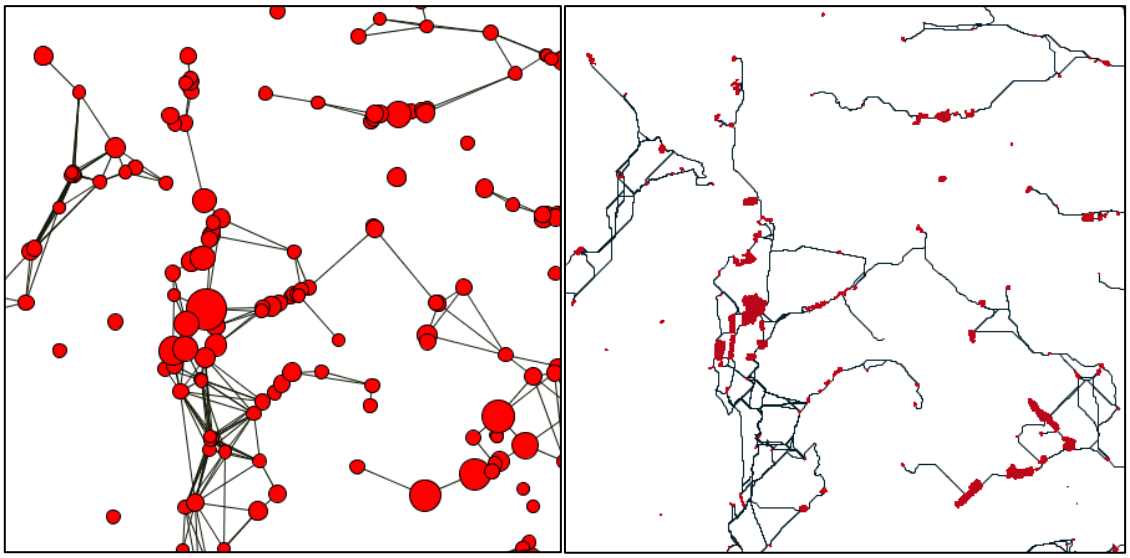
Dans la nouvelle fenêtre qui s'affiche, l'utilisateur retrouve les paramètres du modèle tels qu'ils ont été décrits précédemment. La **résolution** de la grille (en mètres) indique le niveau de précision spatiale selon lequel cette extrapolation est effectuée. Ce critère a une forte incidence sur le temps de calcul nécessaire pour obtenir le résultat. Le résultat est archivé sous forme de couche raster au format *.tif et affiché dans la fenêtre principale.

7. Affichage

7.1. Propriétés des graphes

Les propriétés des graphes sont accessibles par un clic droit sur le nom du graphe. Deux modes de visualisation des graphes sont disponibles :

- La **vue topologique** affiche une vue simplifiée du graphe où les nœuds sont représentés par des cercles et les liens par des lignes droites de centroïde à centroïde.
- La **vue réaliste** affiche les taches d'habitat selon leurs limites réelles et les liens sont représentés par les chemins de moindre coût entre deux taches.



Vue topologique Vue réaliste

Le bouton **Supprimer** permet de supprimer le graphe sélectionné.
Les **Propriétés** rappellent les paramètres utilisés pour la construction du graphe : le nom du graphe, le jeu de liens sélectionné, le type de graphe avec éventuellement le seuil de distance utilisé et le nombre de liens.
Le bouton **Matrice OD** (matrice Origine-Destination) permet de créer un tableau indiquant la distance entre chaque paire de nœuds pour le graphe sélectionné. L'unité de la distance dépend du type de distance utilisé pour ce graphe. L'absence de connexion entre deux nœuds est notée NaN. Cette matrice est archivée dans le dossier du projet sous la forme d'un fichier texte nommé « nom du graphe-odmatrix.txt ».

7.2. Propriétés des objets

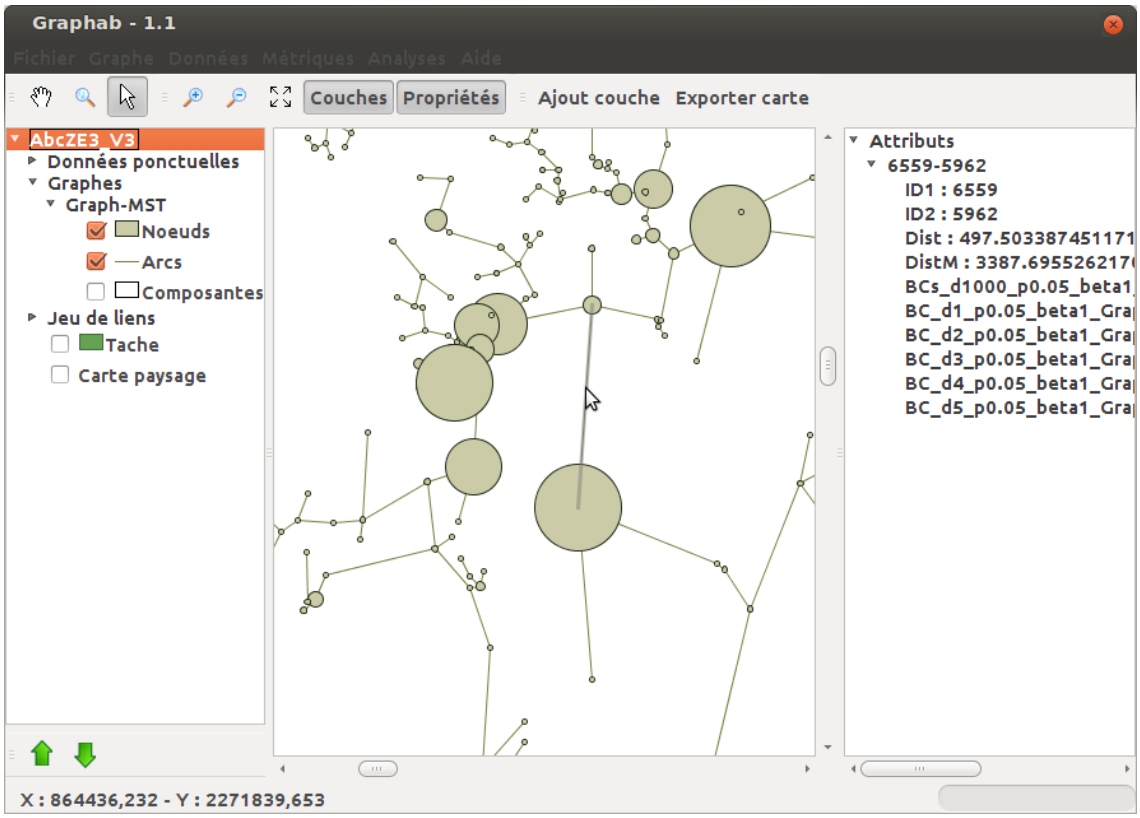
Les propriétés de chaque jeu de liens, des éléments d'un graphe (nœuds, liens ou composantes) et des données ponctuelles sont accessibles par un clic droit sur chacun d'entre eux.

Le menu **Style** donne accès aux paramètres d'affichage des objets : couleur, épaisseur des contours, étiquettes, taille des symboles (pour les nœuds uniquement). Les objets peuvent être représentés de la même manière (symbole unique) ou en fonction d'un attribut. Il est possible d'appliquer une méthode de discrétisation pour classer les objets en fonction des valeurs de l'attribut sélectionné. Par défaut, la légende des objets est affichée dans la table des matières. Il est possible de la cacher en décochant le bouton **Légende**.

Le menu **Export** permet d'exporter les objets dans un fichier vecteur (*.shp) ou un fichier texte (*.txt).
Le menu **Statistiques** permet d'afficher la distribution des valeurs d'un ou plusieurs attributs :

- nuage de point : affiche un nuage de point mettant en relation les valeurs de deux attributs
- histogramme : affiche l'histogramme des valeurs d'un attribut.

Il est également possible d'afficher les valeurs d'un objet particulier en le sélectionnant à l'aide de la flèche. Après sélection, la valeur des attributs s'affiche dans une nouvelle colonne à droite. Il est possible de fermer cette colonne en cliquant sur le menu **Propriétés** dans la barre du haut.



8. Limitations

8.1. Capacités de traitement et limitations

Les méthodes basées sur les graphes sont connues pour fournir un cadre de modélisation efficace, mais peuvent poser une question de capacité de calcul. Deux points spécifiques ont reçu une attention particulière dans Graphab : (1) le calcul des jeux de liens, (2) le calcul des métriques de connectivité. Tous ces calculs ont été optimisés par parallélisation. Ce mode de développement améliore l'efficacité de calcul en utilisant une architecture multiprocesseur, un processeur quatre-cœur permettant théoriquement un calcul quatre fois plus rapide.

Dans la publication suivante [Foltête J.C., Clauzel C., Vuidel G., 2012. A software tool dedicated to the modelling of landscape networks, Environmental Modelling & Software, 38: 316-327], plusieurs tests ont été effectués afin de mesurer la capacité de calcul de Graphab dans différentes configurations. Trois configurations ont été comparées : 1) un cœur (3Go RAM) correspondant à un ordinateur standard ; 2) quatre cœurs (6 Go RAM) correspondant à une station de travail ; 3) 20 cœurs (15 Go RAM) correspondant à un serveur. La carte d'occupation du sol était un fichier raster de 14000*18000 pixels (252 millions de pixels) représentant l'occupation du sol de la région Franche-Comté (France) à une résolution spatiale de 10 mètres. La carte contenait 22 634 taches d'habitats. Le tableau 3 donne quelques temps de calcul pour la création de liens.

Topologie	Distance	Ordinateur standard	Station de travail	Serveur
Complète	Euclidienne	1927s (32 min)	516s (8 min)	133s (2 min)
	Moindre coût	19252s (5h 21 min)	4301s (1h 11 min)	1037s (17 min)
Planaire	Euclidienne	43s	12s	2.6s
	Moindre coût	1080s (18 min)	295s (5 min)	82s (1 min)

Tableau 3. Temps de calcul (en seconde) pour la création de différents jeux de liens

Dans la version 1.1, le temps de calcul des métriques avec les distances intra-taches a été optimisé. La différence de temps de calcul avec ou sans distance intra-tache est maintenant négligeable.

La mémoire allouée au programme joue aussi un rôle important. Si il n'y a pas assez de mémoire vive, le calcul sera plus lent voire même ne pourra pas aboutir (message OutOfMemoryError ou GC Overhead). Le menu **Fichier / Préférences / Mémoire** permet de régler la mémoire allouée pour Graphab. Si vous avez une version 32 bits de Java, Graphab sera limité à environ 2 Go (2000 Mo) de mémoire. Si votre ordinateur a plus que 2 Go de mémoire vive il est fortement conseillé d'installer une version 64 bits de Java pour utiliser la mémoire disponible au-delà des 2 Go.

9. Annexes

9.1. Détail de calcul des métriques

Tableau récapitulatif des métriques présentes dans Graphab 1.1

Famille	Métrique de connectivité	Code	Niveau de calcul			Delta métriques
			Global	Composante	Local	
Métriques pondérées	Flux	F	×	×	×	
	Probabilité de connectivité	PC	×	×		×
	Probabilité de connectivité flux	FPC			×	
	Fractions de delta Probabilité de connectivité	dPC				×
	Indice de centralité intermédiaire	BC			×	
	Indice intégral de connectivité	IIC	×	×		×
Métriques de surface	Taille moyenne des composantes	MSC	×			
	Taille de la plus grande composante	SLC	×			
	Probabilité de coïncidence de classe	CCP	×			
	Taille de composante attendue	ECS	×			
Métriques topologiques	Degré	Dg			×	
	Coefficient de groupement	CC			×	
	Centralité de proximité	CCe			×	
	Excentricité	Ec			×	
	Corrélation de connectivité	CCor			×	
	Nombre de composantes	NC	×			
	Diamètre	GD	×	×		×
	Indice d'Harary	H	×	×		×

Termes mathématiques utilisés

Terme	Signification
n	Nombre de taches
nc	Nombre de composantes
n_k	Nombre de taches de la composante *k*
N_i	Ensemble des taches voisines de la tache *i*
a_i	Capacité de la tache *i* (en général sa surface)
ac_k	Capacité de la composante *k* (somme des capacités des taches qui la composent)
A	Surface de la zone d'étude
d_{ij}	distance entre les taches *i* et *j* (en général la distance de moindre coût qui les sépare)
$e^{-\alpha d_{ij}}$	Probabilité de déplacement entre les taches *i* et *j*
α	Frein de la distance vis-à-vis des déplacements
β	Exposant permettant de pondérer plus ou moins la capacité

Métriques pondérées

Flux (F)		
	Formule	Signification
Niveau global	$F = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} a_j^{\beta} \, e^{-\alpha d_{ij}}$	Pour le graphe entier : somme des dispersions potentielles à partir de toutes les taches.
Niveau local	$F_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_j^{\beta} \, e^{-\alpha d_{ij}}$	Pour la tache focale *i* : somme des capacités des taches différentes de *i* et pondérées en fonction de leur distance minimale à la tache focale en parcourant le graphe. Cette somme est un indicateur de la dispersion potentielle à partir de la tache *i*, ou inversement à destination de la tache *i*.
Valeurs	Les valeurs dépendent de la définition de *a*. Si *a* représente une surface, *F* exprime une surface. Valeur minimale : 0 Valeur maximale : surface totale de l'habitat	
Commentaire	Le chemin du graphe utilisé est celui qui maximise $e^{-\alpha d}$, c'est-à-dire celui qui minimise la distance *d* (ou le coût) entre les taches *i* et *j*. Cette métrique est appelée Area Weigthed Flux (AWF) dans certaines références. Dans Graphab cependant, *a* est plus général car il représente la capacité des taches, qui peut être leur surface ou un autre critère choisi par l'utilisateur. De même la pondération est variable en fonction du paramètre *β*. Dans CS22, AWF est calculé seulement à partir des taches reliées directement à la tache focale, alors que Graphab tient compte des taches indirectement connectées.	
Références	Urban et Keitt, 2001 Saura et Torné, 2009 Foltête et al., 2012a	

Probabilité de Connectivité (PC)		
	Formule	Signification
Niveau global Niveau composante Delta	$PC = \frac{1}{A^2} \sum_{i=1}^{n} \sum_{j=1}^{n} a_i^{\beta} \, a_j^{\beta} e^{-\alpha d_{ij}}$	Pour le graphe entier : somme des produits de la capacité de tous les couples de taches pondérées par leur probabilité d'interaction, divisée par le carré de la zone d'étude. Ce rapport équivaut à la probabilité que deux individus tirés au hasard dans la zone d'étude parviennent à entrer en contact.
Valeurs	Les valeurs correspondant à une probabilité. Valeur minimale : 0 Valeur maximale : 1	
Commentaire	Pour chaque couple de taches, le chemin du graphe utilisé est celui qui maximise $e^{-\alpha d}$, c'est-à-dire celui qui minimise la distance *d* (ou le coût) entre les taches *i* et *j*. Dans CS22, la pondération des capacités est fixée à 1 ; dans Graphab elle peut être modifiée. Si *a* ne représente pas la surface des taches, le résultat n'est plus une probabilité.	
Référence	Saura et Pascual-Hortal, 2007	

Probabilité de Connectivité Flux (FPC)		
	Formule	Signification
Niveau local	$FPC_i = \frac{1}{A^2}\sum_{j=1}^{n} a_i^{\beta} a_j^{\beta} e^{-\alpha d_{ij}}$	Somme des produits de la capacité de la tache focale avec toutes les autres taches, pondérées par leur probabilité d'interaction, divisée par le carré de la surface de la zone d'étude.
Valeurs	Valeur minimale : 0 Valeur maximale : 1	
Commentaire	Pour chaque couple de taches, le chemin du graphe utilisé est celui qui maximise $e^{-\alpha d}$, c'est-à-dire celui qui minimise la distance *d* (ou le coût) entre les taches *i* et *j*. Cette métrique est simplement la contribution locale d'une tache à l'indice *PC*, puisque $PC = \sum_i FPC_i$. Elle équivaut à l'indice dPC_{flux} non divisé par la valeur globale du PC. Cependant la métrique FPC est obtenue plus rapidement que dPC_{flux} car son calcul n'est pas fondé sur la suppression itérative des taches (mode delta).	

Fractions de delta Probabilité de Connectivité (dPC, dPC_{area}, dPC_{flux}, $dPC_{connector}$)		
	Formules	Signification
Delta	$dPC_i = \frac{(PC - PC_i')}{PC}$ $dPC_i = dPC_{area} + dPC_{flux} + dPC_{connector}$ $dPC_{area} = \frac{a_i^{2\beta}}{A^2 PC}$ $dPC_{flux} = \frac{FPC_i}{PC}$	Taux de variation entre la valeur de l'indice *PC* et la valeur PC' représentant la suppression de la tache *i*. La valeur de *dPC* se décompose en trois parties : - dPC_{area} correspond à la variation due à la perte de la surface suite à la suppression de la tache *i* ; - dPC_{flux} correspond à la variation due à la perte des interactions entre la tache *i* et les autres taches ; - $dPC_{connector}$ correspond à la variation due à la modification des chemins reliant les autres taches et passant initialement par *i*.
Valeurs	Valeur minimale : 0 Valeur maximale : 1	
Commentaire	Dans CS22, la pondération des capacités est fixée à 1 ; dans Graphab elle peut être modifiée. Si *a* ne représente pas la surface des taches, le résultat n'est plus une probabilité et dPC_{area} n'exprime pas une perte de surface mais de capacité.	
Référence	Saura et Rubio, 2010	

Indice de centralité intermédiaire (BC)		
	Formule	Signification
Niveau local	$BC_i = \sum_j \sum_k a_j^{\beta}\, a_k^{\beta} e^{-\alpha d_{jk}}$ $j,k \in \{1..n\}, k < j, i \in P_{jk}$	Somme des plus courts chemins passant par la tache focale *i*, chaque chemin étant pondéré par le produit des capacités des taches reliées et de leur probabilité d'interaction. P_{jk} représente l'ensemble des taches traversées par le plus court chemin entre les taches *j* et *k*.
Valeurs	Les valeurs dépendent du paramétrage. Elles correspondent à un poids de transit potentiel. Valeur minimale : 0 Valeur maximale : surface totale de l'habitat au carré.	
Commentaire	Avec un paramétrage de α à 0 et β à 0 (pondération uniforme des chemins), on retrouve l'indice BC tel qu'il est utilisé dans d'autres types de graphes. Un paramétrage de α=1 et β=0 donne aux chemins un poids proportionnel au produit des capacités des taches qu'ils relient, quelle que soit leur distance. Dans Foltête et al. (2012a ; 2012b), l'indice BCl a été proposé pour pondérer plus fortement les chemins supérieurs au critère choisi (distance de dispersion par exemple). Mais des tests ont ensuite montré que cet indice était très fortement corrélé au BC pondéré avec β=0. Dans Bodin et Saura (2010), le BC_{pc} correspond au BC pondéré avec *d* égal à la distance de dispersion, α tel que $e^{-\alpha d} = 0.05$ et $\beta = 1$.	
Références	Bodin et Saura, 2010 Foltête et al., 2012b	

Indice Intégral de Connectivité (IIC)		
	Formule	Signification
Niveau global Niveau composante Delta	$IIC = \frac{1}{A^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{a_i a_j}{1 + nl_{ij}}$	Pour le graphe entier : produit des capacités des taches divisé par le nombre de liens qui les sépare, la somme étant divisée par le carré de la surface de la zone d'étude. IIC est construit comme l'indice PC, mais en utilisant l'inverse d'une distance topologique plutôt qu'une fonction exponentielle négative de la distance basée sur l'impédance des liens.
Valeurs	Valeur minimale : 0 Valeur maximale : 1	
Référence	Pascual-Hortal et Saura, 2006	

Métriques de surface

Taille moyenne des composantes (MSC)		
	Formule	Signification
Niveau global	$MSC = \frac{1}{nc}\sum_{k=1}^{nc} ac_k$	Pour le graphe entier : moyenne des capacités des composantes.
Valeurs	Valeur minimale : capacité minimale Valeur maximale : SLC	

Taille de la plus grande composante (SLC)		
	Formule	Signification
Niveau global	$SLC = \max\{ac_k\}$	Pour le graphe entier : plus grande capacité des composantes.
Valeurs	Valeur minimale : capacité minimale Valeur maximale : capacité maximale	

Probabilité de coïncidence de classe (CCP)		
	Formule	Signification
Niveau global	$CCP = \sum_{k=1}^{nc}\left(\frac{ac_k}{\sum_l ac_l}\right)^2$	Pour le graphe entier : probabilité que deux points choisis au hasard sur le graphe appartiennent à la même composante.
Valeurs	Valeur minimale : capacité minimale (autant de composantes que de taches et capacités uniformes) Valeur maximale : somme des capacités (1 seule composante)	
Référence	Pascual-Hortal et Saura, 2006	

Taille de composante attendue (ECS)		
	Formule	Signification
Niveau global	$ECS = \frac{1}{\sum_k ac_k}\sum_{k=1}^{nc} ac_k^2$	Pour le graphe entier : ?
Valeurs	Valeur minimale : capacité minimale (autant de composantes que de taches et capacités uniformes) Valeur maximale : somme des capacités (1 seule composante)	
Référence	O'Brien et al., 2006	

Métriques topologiques

Indice d'Harary (H)		
	Formule	Signification
Niveau global Niveau composante	$H = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{nl_{ij}}$	Somme de l'inverse du nombre de liens séparant tous les couples de taches.
Valeurs	Valeur minimale : capacité minimale (autant de composantes que de taches et capacités uniformes) Valeur maximale : somme des capacités (1 seule composante)	
Commentaire	Pour les couples de taches non connectées par un chemin on a : $nl_{ij} = +\infty$	
Référence	Ricotta et al., 2000	

Diamètre (GD)		
	Formule	Signification
Niveau global Niveau composante Delta	$GD = \max_{ij} d_{ij}$ $GD = \max_{i} Ec_i$	Distance la plus grande entre deux taches du graphe
Valeurs	Valeur minimale : 0 Valeur maximale : $+\infty$	
Commentaire	Quand les nœuds *i* et *j* ne sont pas connexes $d_{ij} = 0$ Cette métrique est la version globale de la métrique Ec_i	

Nombre de composantes (NC)		
	Formule	Signification
Niveau global	$NC = nc$	Nombre de composantes du graphe
Valeurs	Valeur minimale : 1 Valeur maximale : n	

Degré du noeud (Dg)		
	Formule	Signification
Niveau local	$Dg_i = \lvert N_i \rvert$	Nombre de taches voisines de la tache *i*
Valeurs	Valeur minimale : 0 Valeur maximale : n	
Commentaire	Il y a équivalence entre le degré d'un nœud et le nombre de nœuds voisins car les graphes ne sont pas orientés et ne contiennent pas de boucle.	

Coefficient de groupement (CC)		
	Formule	Signification
Niveau local	$CC_i = \frac{1}{\lvert N_i \rvert(\lvert N_i \rvert - 1)} \sum_{j \in N_i} \lvert N_i \cap N_j \rvert$	Rapport du nombre de nœuds voisins de *i* voisins entre eux sur le total possible.
Valeurs	Valeur minimale : 0 Valeur maximale : 1	
Commentaire	Si $\lvert N_i \rvert \leq 1 \rightarrow CC_i = 0$	

Centralité de proximité (CCe)		
	Formule	Signification
Niveau local	$CCe_i = \frac{1}{n_k - 1} \sum_{\substack{j=1 \\ j \neq i}}^{n_k} d_{ij}$	Distance moyenne de la tache *i* vers toutes les autres taches de sa composante *k*.
Valeurs	Valeur minimale : 0 Valeur maximale : $+\infty$	
Commentaire	Si $n_k = 1 \rightarrow CCe_i = 0$	

Excentricité (Ec)		
	Formule	Signification
Niveau local	$Ec_i = \max_j d_{ij}$	Distance maximale de la tache *i* vers une autre tache de sa composante.
Valeurs	Valeur minimale : 0 Valeur maximale : $+\infty$	

Corrélation de connectivité (CCor)		
	Formule	Signification
Niveau local	$CCor_i = \frac{\lvert N_i \rvert^2}{\sum_{j \in N_i} \lvert N_j \rvert}$	Rapport entre le degré du nœud *i* et le degré de ses nœuds voisins *j*
Valeurs	Valeur minimale : 0 Valeur maximale : $\lvert N_i \rvert$	
Commentaire	Si $\lvert N_i \rvert = 0 \rightarrow CCor_i = 0$	

9.2. Références bibliographiques

Bodin O., Saura S., 2010. Ranking individual habitat patches as connectivity providers: Integrating network analysis and patch removal experiments. Ecological Modelling 221, 2393–2405.

Foltête J.C., Clauzel C., Tournant P., Vuidel G., 2012a, Integrating graph-based connectivity metrics into species distribution models, Landscape Ecology 27, 557-569.

Foltête J.C., Clauzel C., Vuidel G., 2012. A software tool dedicated to the modelling of landscape networks. Environmental Modelling and Software 38, 316-327.

Minor E.S., Urban D.L., 2008. A graph-theory framework for evaluating landscape connectivity and conservation planning. Conservation Biology 22, 297–307.

O'Brien D., Manseau M., Fall A., Fortin M.J., 2006. Testing the importance of spatial configuration of winter habitat for woodland caribou: An application of graph theory. Biological Conservation 130, 70–83.

Pascual-Hortal L., Saura S., 2006. Comparison and development of new graph-based landscape connectivity indices: towards the priorization of habitat patches and corridors for conservation. Landscape Ecology 21, 959–967

Rayfield B., Fortin M.J., Fall A., 2011. Connectivity for conservation: a framework to classify network measures. Ecology 92, 847–858.

Ricotta C., Stanisci A., Avena G.C., Blasi C., 2000, Quantifying the network connectivity of landscape mosaics: a graph-theoretical approach, Community Ecology 1(1), 89–94.

Saura S., Pascual-Hortal L., 2007. A new habitat availability index to integrate connectivity in landscape conservation planning: Comparison with existing indices and application to a case study. Landscape and Urban Planning 83, 91–103.

Saura S., Torné J., 2009. ConeforSensinode 2.2: A software package for quantifying the importance of habitat patches for landscape connectivity. Environmental Modelling and Software 24, 135–139.

Saura S., Rubio L., 2010. A common currency for the different ways in which patches and links can contribute to habitat availability and connectivity in the landscape. Ecography 33, 523–537.

Urban D.L., Keitt T.H., 2001. Landscape connectivity: a graph theoretic approach. Ecology 82, 1205–1218.